

Introduction to AI Policies, Guidelines, and Frameworks

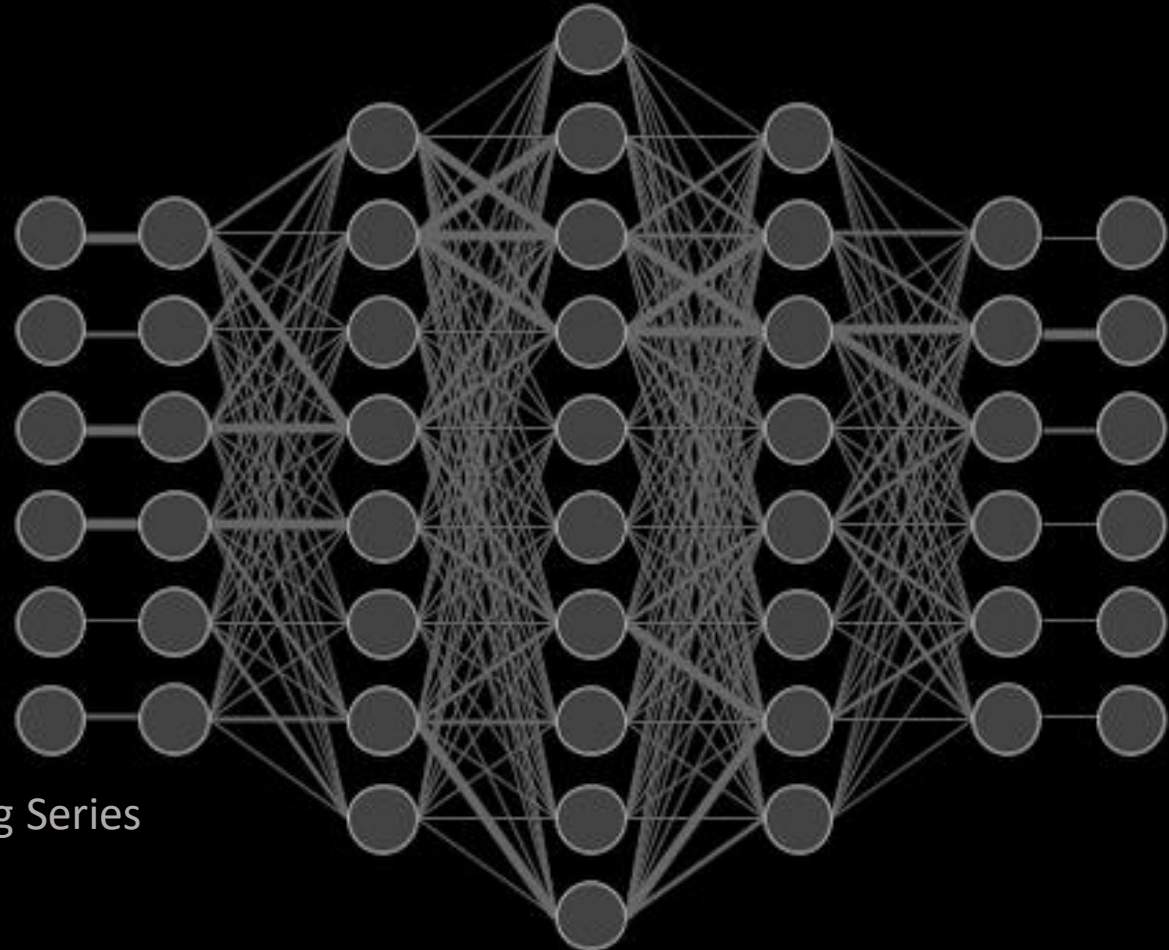


Stephen Downes

November 16, 2023

Maskwacis Cultural College Online Microlearning Series

<https://www.downes.ca/presentation/575>



Introduction

This workshop will look at recent policies, guidelines and frameworks developed by governments and institutions for the ethical and responsible use of artificial intelligence.

How are they created, and what things do they address?

You will have the opportunity to identify important values and principles in order to develop a starting point for discussion relevant to your own context.

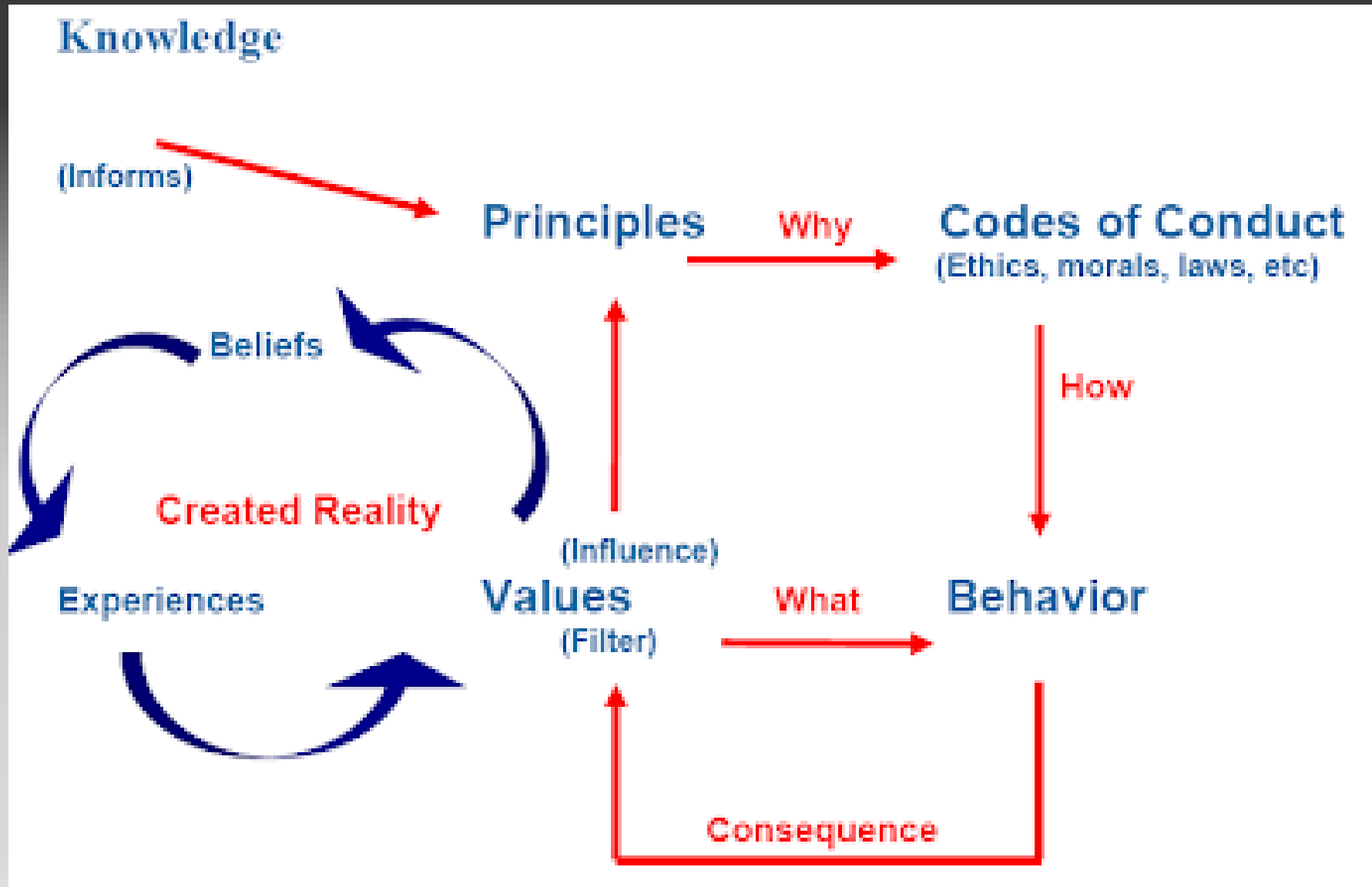
The Question

Artificial Intelligence (AI) is a new type of technology that offers many potential benefits but comes with unknown risks.

How do we as a society govern ourselves with regard to the creation and use of this technology?

Governments, industry and societies have been grappling with this question.





<https://www.mvf-knowledge-base.com/2011/12/whats-difference-between-values-ethics.html>

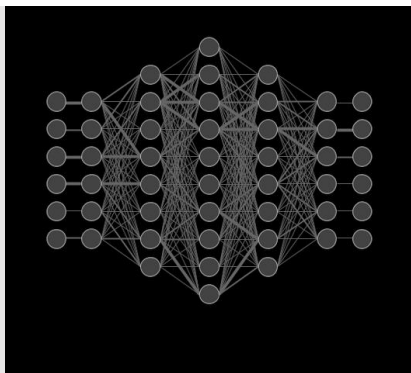
Approaches to Ethics

- **Virtue and character** – “Be all that you can be”
- **Duty** – “with great power comes great responsibility”
- **Consequentialism** - ranging from “do no harm” to “the ends justify the means”.
- **Social contract** – “We the people, in order to form a more perfect union...”



<https://ethics.mooc.ca/cgi-bin/page.cgi?module=9>

What Are Your Values?



- Scan the QR code to the left or go to the following URL (probably works much better on a wider screen)

<http://www.downes.ca/post/75794/rd>

- Enter your votes for the first ‘election’ (we’ll do the next two later in the session).

An Ethical Codes Reader

Survey of 70 ethical codes of conduct in professions dealing with knowledge and technology

- Brief summary of each code
- Link to the code web page
- Text of relevant sections of the code



<https://docs.google.com/document/d/1mv9VxblyGvBaFwHSvtyN1a3iwAKws6SZDGuDTIYMZ48/edit#heading=h.roink2k5jm7f>

Image: <https://www.aje.com/services/research-promotion/>

Some Major Statements

House of Lords Select Committee on AI

- It would be better to focus on “the consequences of AI, rather than on the way it works”, in a way that empowered individuals to exercise their rights (para 51).
- Four of the five principles in fact represent a consequentialist approach, however one principle states that AI “should operate on principles of intelligibility and fairness” (para 417).

<https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>

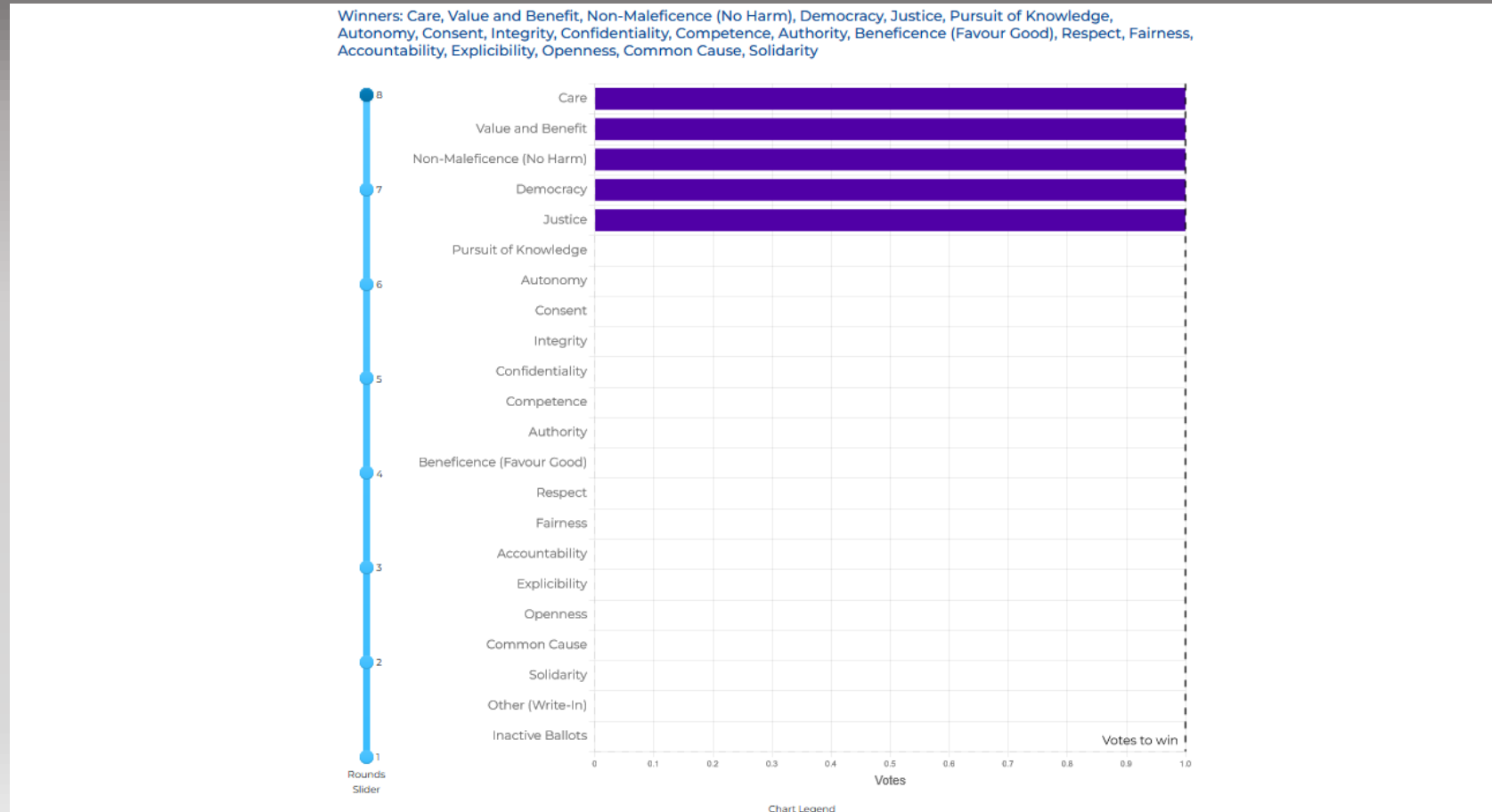
Some Major Statements

US Whitehouse - Blueprint for an AI Bill of Rights

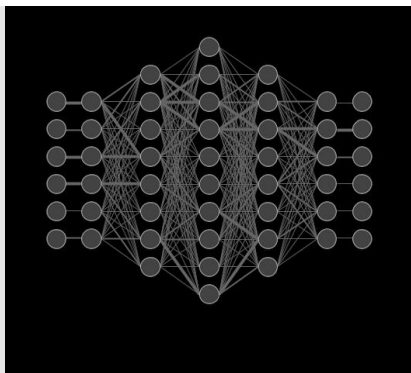
- Safe and Effective Systems
- Algorithmic Discrimination Protections
- Data Privacy
- Notice and Explanation
- Human Alternatives, Consideration, and Fallback

<https://www.whitehouse.gov/ostp/ai-bill-of-rights/>

Is There a Consensus?



Who Are We Responsible To?



- Scan the QR code to the left or go to the following URL (probably works much better on a wider screen)

<http://www.downes.ca/post/75794/rd>

- Enter your votes for the **second** 'election' (we'll do the third later in the session).

Some Major Statements

TBC Canada - Directive on Automated Decision-Making

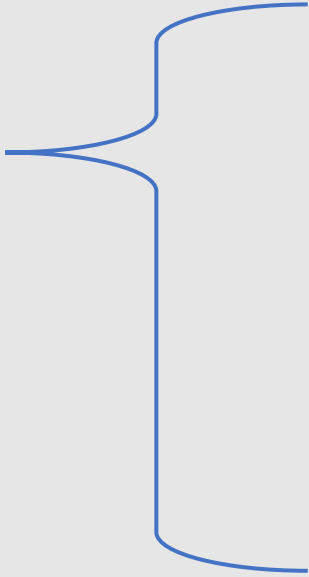
- Algorithmic Impact Assessment

- Transparency

- Quality Assurance

- Recourse

- Reporting

- 
- Testing and monitoring outcomes
 - Data quality
 - Data governance
 - Peer review
 - Gender-based analysis Plus
 - Employee Training
 - IT and business continuity management
 - Security
 - Legal
 - Ensuring human intervention

<https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

Some Major Statements

Amnesty - Equality and Non-discrimination

Using the framework of international human rights law

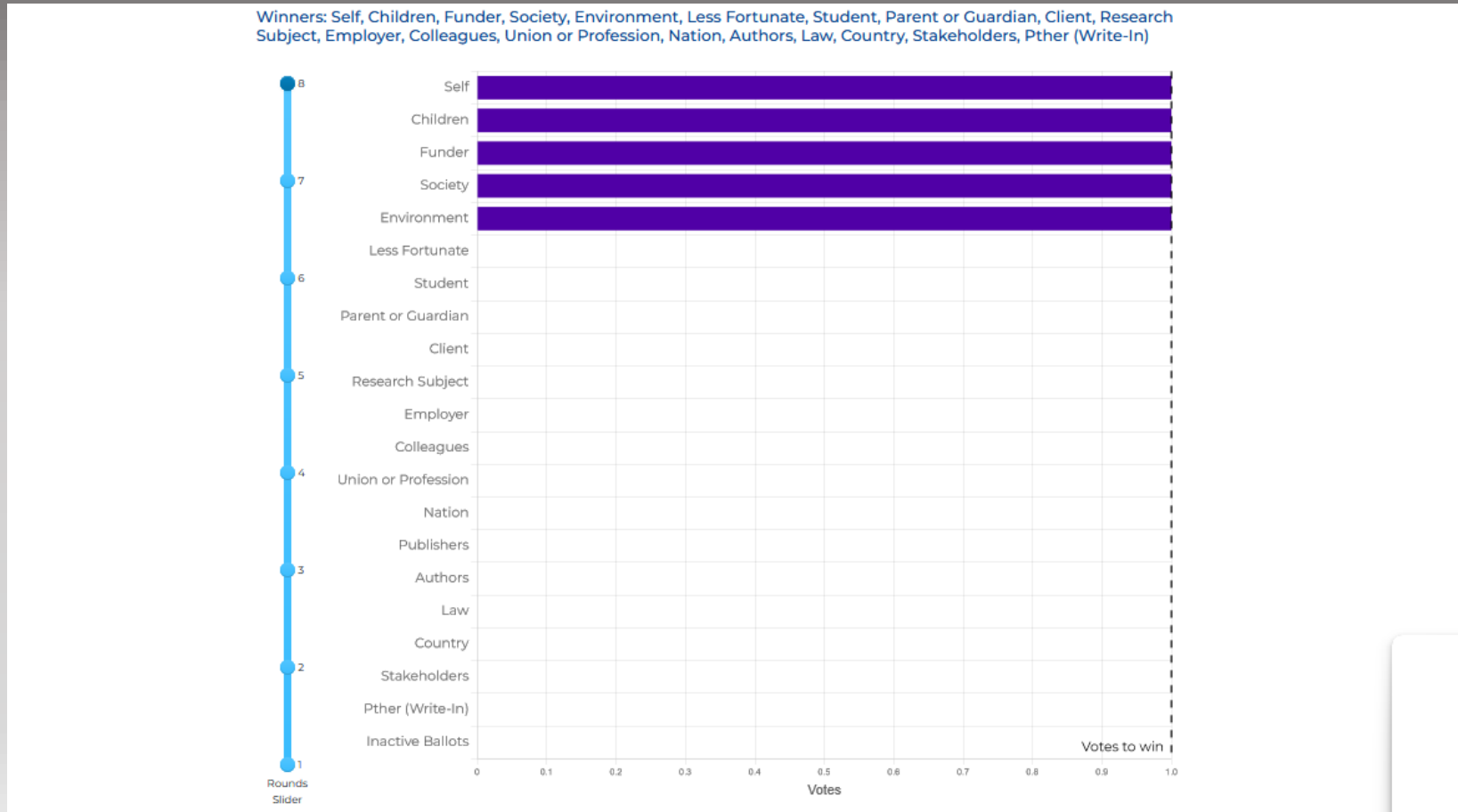
- The right to equality and non-discrimination
- Preventing discrimination
- Protecting the rights of all individuals and groups: promoting diversity and inclusion

Duties of states: human rights obligations

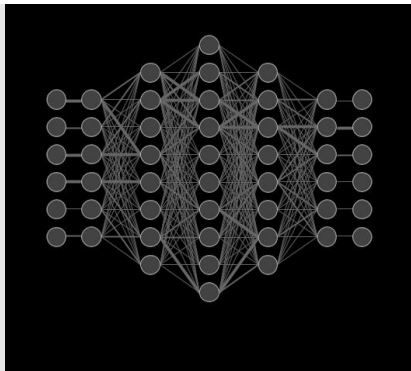
- Identify risks
- Ensure transparency and accountability
- Enforce oversight

https://www.accessnow.org/wp-content/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf

Who Matters?



Foundational Principles



- Scan the QR code to the left or go to the following URL (probably works much better on a wider screen)

<http://www.downes.ca/post/75794/rd>

- Enter your votes for the **third** 'election'

Some Major Statements

Canada - Guide on the use of Generative AI

FASTER Principles:

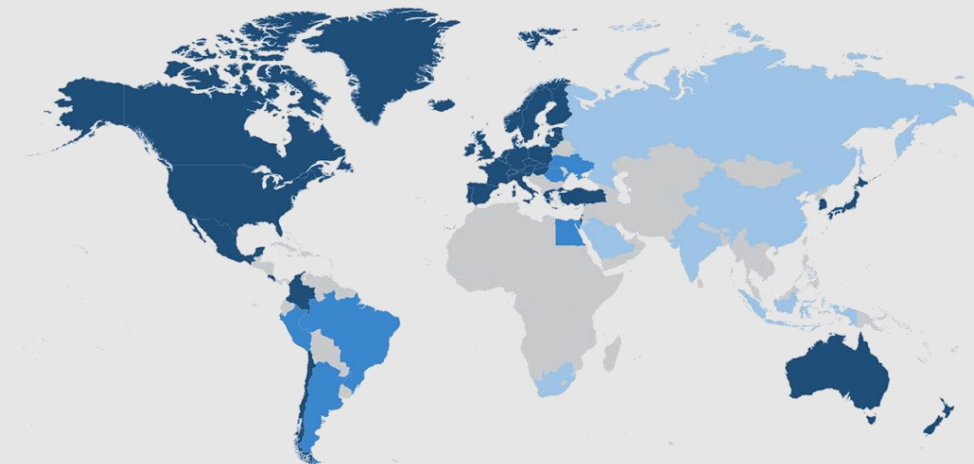
- **Fair:** do not include or amplify biases, comply with human rights, accessibility, and procedural and substantive fairness obligations
- **Accountable:** making sure it is factual, legal, ethical, and compliant with the terms of use
- **Secure:** security classification of the information, that privacy and personal information are protected
- **Transparent:** identify content that has been produced using generative AI and notify users
- **Educated:** learn about the strengths, limitations and responsible use of the tools
- **Relevant:** identify appropriate tools for the task; AI tools aren't the best choice in every situation

<https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html>

Some Major Statements

OECD AI Principles overview

- Inclusive growth, sustainable development and well-being
- Human-centred values and fairness
- Transparency and explainability
- Robustness, security and safety
- Accountability



<https://oecd.ai/en/ai-principles>

■ OECD members ■ G20 principles, based on OECD
■ Adherents *Singapore is an adherent

Some Major Statements

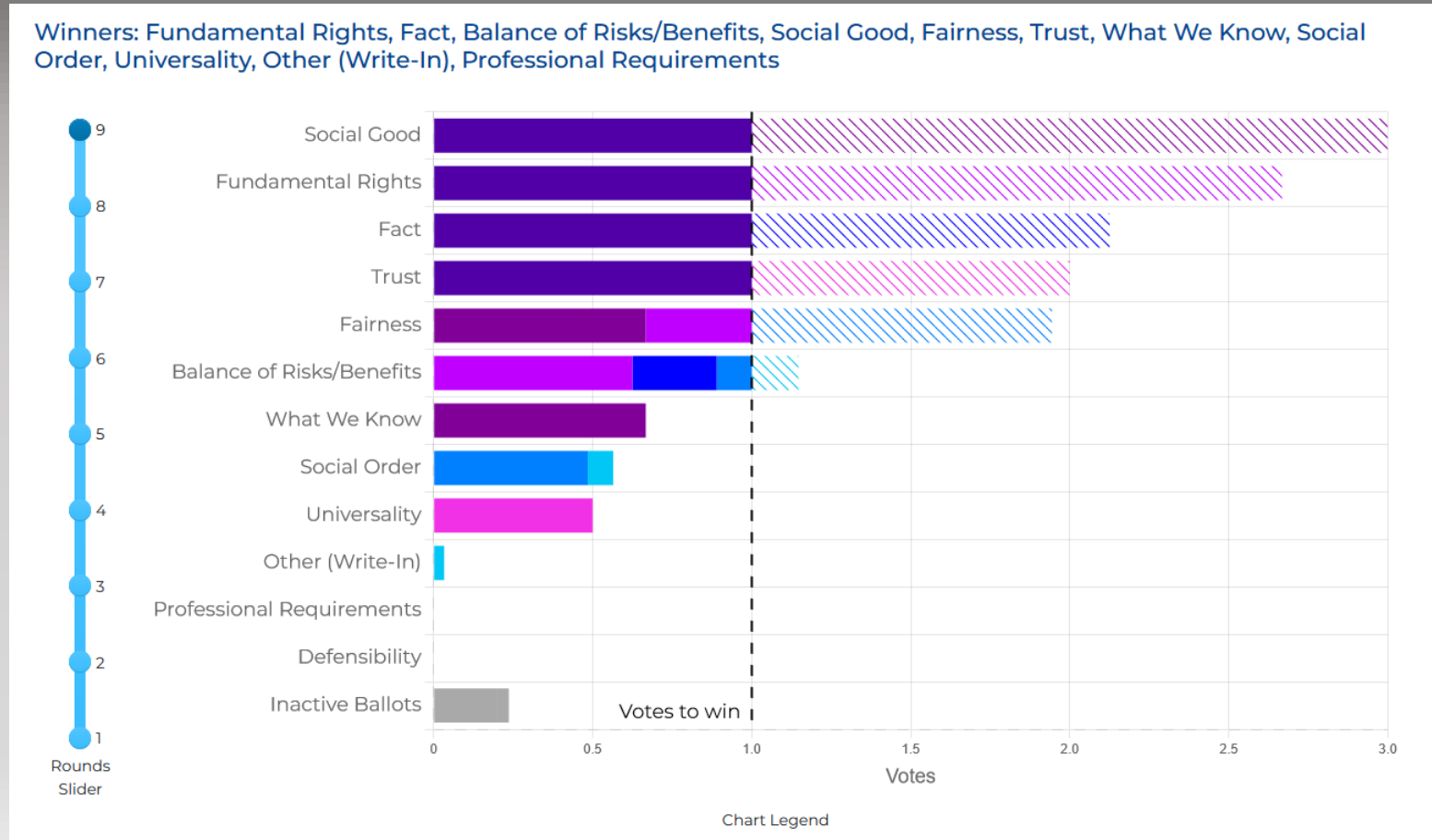
University of Helsinki – AI Ethics MOOC

The five principles of AI ethics answer different questions and focus on different values:

- Should we use AI for good and not for causing harm? (the principle of beneficence/ non-maleficence)
- Who should be blamed when AI causes harm? (the principle of accountability)
- Should we understand what, and why AI does whatever it does? (the principle of transparency)
- Should AI be fair or non-discriminative? (the principle of fairness)
- Should AI respect and promote human rights? (the principle of respecting basic human rights)

<https://ethics-of-ai.mooc.fi/>

What Matters?



Your Questions

Some more resources:

<https://bryanalexander.org/classes-and-teaching/what-should-my-grad-students-read-about-emerging-ai/>

Stephen Downes

<https://www.downes.ca>

